

Post Peer Review version.

Mercedes Durham et al. (2012). Constant Linguistic Effects in the Diffusion of *be like*.

*Journal of English Linguistics* 40(4) 316–337. DOI: 10.1177/0075424211431266

## Constant Linguistic Effects in the Diffusion of *be like*

Mercedes Durham<sup>1</sup>, Bill Haddican<sup>2</sup>, Eytan Zweig<sup>3</sup>, Daniel Ezra Johnson<sup>4</sup>, Zipporah Baker<sup>3</sup>,

David Cockeram<sup>3</sup>, Esther Danks<sup>3</sup>, Louise Tyler<sup>3</sup>

1 Cardiff University, 2 Queens College Cuny, 3 University of York, 4 Lancaster University

### Abstract

This paper examines change in social and linguistic effects on *be like* usage and acceptability. Results from two studies are presented. The first set of data comes from a trend study with samples of U.K. University undergraduates collected in 1996 and 2006. While the effect of subject person, morphological tense, and quote content is constant in our two samples, the effect of speaker sex decreases. The second study is a judgment experiment with 121 native speakers of U.S. English, examining acceptability of *be like* in environments biasing direct speech and reported thought readings. The analysis reveals no interaction between age and the reported thought/direct speech contrast, suggesting no support for change in this effect on *be like* acceptability in apparent time. The two studies therefore converge in suggesting no evidence of change in linguistic constraints on *be like* as it has diffused into U.K. and U.S. Englishes.

**Keywords:** *quotatives; be like; British English; real time; sociolinguistic; language variation and change*

## Introduction

The spread of *be like* as a quote introducer has received much attention over the past fifteen years and has provided an unprecedented scope for linguists to examine extremely rapid change across generations (Buchstaller 2006a, 2006b; Buchstaller & D’Arcy 2009; Cukor-Avila 2002; Dailey-O’Cain 2000; Ferrara & Bell 1995; Macaulay 2001; Romaine & Lange 1991; Tagliamonte & D’Arcy 2004, 2007; Tagliamonte & Hudson 1999). As much recent literature has noted, the process of spread of *be like* now has sufficient time depth to allow for detailed comparisons across age groups (Buchstaller 2006a; Cukor-Avila 2002; Ferrara & Bell 1995; Tagliamonte & D’Arcy 2004, 2007). Much of the recent literature on *be like* has therefore focused on whether some well known constraints on English quote introducers have changed through the course of diffusion of *be like* in various dialects. The most careful and detailed studies of this sort have been Tagliamonte and D’Arcy’s (2004, 2007) studies of change in Toronto English, which indicate ongoing change in social and linguistic constraints on *be like* as the form enters the local grammar. Most notably, among later age cohorts, Tagliamonte and D’Arcy find an increase in the speaker sex effect with women favoring *be like*, and a weakening of the effect of quote content, where internal dialogue (reported thought) contexts favor *be like* over direct speech contexts. Tagliamonte and Hudson (1999:167-169) and Tagliamonte and D’Arcy (2004:511, 2007:202) suggest that these patterns represent general pathways of grammaticalization of *be like*; that is, it’s a change which, *ceteris paribus*, all dialects into which *be like* is diffusing will undergo.

The goal of this article is to assess evidence for similar changes in constraints in varieties of English where quotatives have been less studied. In particular, this article considers two sets of data—one with U.K. English speakers and a second with U.S. English speakers—and tests whether the interaction between age group and other well known social and linguistic effects on *be like* is more general in nature. The first set comes from a trend study comparing

quotative usage in a 2006 corpus of speech from University of York undergraduate students aged 18-22 with those from a similarly constructed sample from 1996. This comparison allows us to assess possible cross-generational change in constraints on *be like* use in York. The second set of data comes from a judgment experiment, conducted in 2009 with speakers of American English, comparing acceptability of *be like* and *say* in contexts biasing direct speech and reported thought readings.

Our results suggest two main findings. First, the corpus study shows a significant interaction between sex and age group with a smaller sex effect in the later data set. These results are in keeping with previous studies suggesting mutability of social factors across age groups and dialects as *be like* expands (Buchstaller & D’Arcy 2009; Ferrara & Bell 1995; Tagliamonte & D’Arcy 2007;). Second, the corpus data and the experimental results both reveal that no interaction between age and the classic linguistic constraints on *be like* discussed in much previous literature. The contrast between our findings and previous corpus findings from other locales therefore supports skepticism with regard to universal tendencies in constraint shifts in grammaticalization of *be like*. Instead our results support findings from the historical syntax literature of a constant rate effect on syntactic change—that is, for any single abstract process of grammatical change, a constancy in the effect of different contexts across generations of speakers advancing the change (Kroch 1989, 1994, 2000).

Our discussion is organized as follows. We begin by discussing previous literature examining change on constraints on *be like* usage. We then present our two original studies focused on change in linguistic effects on *be like*. The final section summarizes some implications of the data presented.

### **The Expansion of *Be Like***

In contemporary English speech there is considerable variation in verbs which can be used to introduce direct speech as in (1)-(3).

- (1) I was like “Easy tiger.”
- (2) She said “Let’s go.”
- (3) He went “Calm it love.”

Over the past two decades an extensive body of literature has focused on the expansion of the *be like* variant in (1). This literature has documented the spread of *be like* in several English-speaking societies globally starting in the U.S. (Cukor-Avila 2002; Ferrara & Bell 1995; Singler 2001), and later in Canada (Tagliamonte & D’Arcy 2004, 2007; Tagliamonte & Hudson 1999), the U.K. (Buchstaller 2006a; Macaulay 2001; Tagliamonte & Hudson 1999), and New Zealand (Buchstaller & D’Arcy 2009).

Following work by Ferrara and Bell (1995) and Tagliamonte and Hudson (1999), much of this literature has focused on continued grammaticalization of *be like* during this process of expansion. In an effort to track these changes, several recent studies have examined changes in the way *be like* use is constrained internally and externally in corpus data. Below we introduce four constraints frequently discussed in this literature.

### *Speaker Sex*

Several studies have reported a speaker sex effect on *be like* usage. The most typical finding in the literature is that women tend toward innovative *be like* more than men (Blyth et al. 1990; Ferrara & Bell 1995; Macaulay 2001; Tagliamonte & D’Arcy 2004; Tagliamonte & Hudson 1999). This pattern is in keeping with evidence from perceptual data. In particular,

Dailey-O’Cain’s (2000) U.S. study and Buchstaller’s (2006a) U.K. study both suggest that quotative *be like* use is associated with young women.

Much of the literature, moreover, indicates that sex effects on quotative use are mutable diachronically and across communities. In a three-year trend study in Texas in the early 1990s, Ferrara and Bell (1995) found evidence of neutralization of the sex effect: in 1990 women used *be like* twice as frequently as men (15 vs. 29 percent); however, in later, similarly constructed samples in 1992 and 1994, men and women used *be like* at roughly equal rates. During this period, the overall rate of *be like* use increased steadily. These data might be taken to indicate that as *be like* diffuses, the effect of speaker sex may weaken. However, Tagliamonte and D’Arcy (2004) describe the opposite pattern in Canada: with increasing use of *be like*, a sex difference emerged between 1995 and 2002/3. Buchstaller and D’Arcy (2009) also find conflicting results for gender across American, British, and New Zealand corpora. Based in part on these results, Tagliamonte and Hudson (1999) and Tagliamonte and D’Arcy (2004) hypothesize the opposite relationship between diffusion of *be like* and sex differentiation: the further *be like* diffuses, “the more likely it is to differentiate male and female speech” (Tagliamonte & Hudson 1999:167).

### *Quote Content*

A second constraint on *be like* usage discussed in recent literature concerns the interpretation of the quoted material. The availability of *be like* in contexts such as (1) appears to have emerged through a reanalysis of sequences of *be* + discourse marker *like*. That is, prior to its emergence as an introducer of direct speech, *be like* could be used to describe states of individuals in contexts with predicate adjectives (4), with non-lexicalized sounds as in (5), and in “internal dialogue” as in (6).

- (4) I was like devastated.
- (5) She was like “Ugh.”
- (6) I was like “Never again.”

In studying the increase of quotative *be like*, several authors have examined its distribution across age cohorts in some of the above contexts. Early studies report that *be like* is disfavored in contexts introducing direct speech and favored before non-lexicalized sounds and internal dialogue (Ferrara & Bell 1995; Macaulay 2001; Tagliamonte & Hudson 1999). More recently, however, Tagliamonte and D’Arcy (2004) show that for one of the three age groups in their 2002/3 sample (17-19 year olds), speakers used *be like* to a greater extent in direct speech than internal dialogue. Similarly, Tagliamonte and D’Arcy (2007) demonstrate that young speakers aged 17-29, showed the weakest effect of quote content in their sample. The authors plausibly interpret these results as evidence of continued grammaticalization of *be like* as an introducer of direct speech.

### *Subject Person*

Results reported in the literature concerning the effect of subject person on the variation have been more consistent. As it emerged as a dialogue introducer, *be like* appears to have been originally favored with first-person subjects, a fact plausibly related to its role as an introducer of internal dialogue. Ferrara and Bell (1995) report that this effect weakens during the period covered by their study, and they interpret this shift as evidence of expansion of function of *be like*. Most other studies, however, have reported much greater constancy in the effect of subject person, with first-person subjects favoring *be like* usage and second- and third-person subjects disfavoring it (Buchstaller & D’Arcy 2009; Cukor-Avila 2002; Tagliamonte & D’Arcy 2004, 2007; Tagliamonte & Hudson 1999).

### *Tense and Aspect*

Finally, several authors have focused on the effect of tense on *be like* usage. In particular, Blyth et al. (1990), Romaine and Lange (1991), and Singler (2001) all report that *be like* is favored in present tense contexts and disfavored in the past tense. More recently, Tagliamonte and D'Arcy (2007) use a three-way coding for tense, distinguishing past, present, and historical present. Their analysis of variation in a corpus of speakers ranging in age from 9-39 in Toronto indicates that, across several age groups, historical present contexts most consistently favored *be like* use, followed by present tense contexts and finally past tense contexts. Buchstaller and D'Arcy (2009) corroborate this pattern for the corpora of American and New Zealand English as well; for their United Kingdom corpus past contexts were most favored. Our study uses a two-way distinction focusing solely on the tense of the verb; the token numbers of present are too low in one of our data sets (12 tokens or 2.4 percent) to examine it separately.

To our knowledge, none of the literature on grammaticalization of *be like* has touched on variation and change in the aspectual behaviour of *be like*. In particular, the claim that *be like* is undergoing a process of reanalysis from stative *be* + discourse *like* to a quotative verb akin to accomplishment verbs such as *say* and *go* suggests the possibility that *be like* will come to behave like the latter in progressive environments. That is, stative verbs like *have*, *know* and *be* are unlike activity and accomplishment predicates in that they are canonically poor in progressives, as illustrated in (7)-(11).

- (7) Tina is eating the sandwich.
- (8) Terry was smoking.
- (9) \*I'm having money.

(10) \*I'm knowing French.

(11) \*Terry is being ill.

If *be like* is indeed coming to behave syntactically like *go* and *say*, then we might expect it to begin to appear in progressive contexts, as in (12)-(14) from the 2006 data set.<sup>1</sup> We examine this possibility below.

(12) The cat is going “Meow.”

(13) She was going “Don’t touch.”

(14) He was saying “Uhm oh we’ve got to move her.”

Tables 1 and 2 summarize the evolution of the above four constraints on *be like* as described in two influential sets of real-time studies of quotatives in North America: Ferrara and Bell’s (1995) data from Texas and Tagliamonte and D’Arcy’s (2004, 2007) data from Toronto.

**TABLE 1**

Ferrara & Bell’s (1995) constraint evolutions for *be like*

<i>Constraint</i>	<i>Early stages</i>	<i>Later stages</i>
Sex	Women > Men	Neutralization
Subject Person	First > third	Expansion into third person
Content	Internal dialogue > direct speech	Expansion into direct speech



**TABLE 2**Tagliamonte and D'Arcy's (2007) constraint evolutions for *be like*

<i>Constraint</i>	<i>Early stages</i>	<i>Later stages</i>
Sex	Women > Men	Increased differentiation
Subject Person	First > third	Constancy of effect
Content	Internal dialogue > direct speech	Expansion into direct speech
Tense	Present tense favored	Historical present favored

Our goals in view of the foregoing literature are twofold. First, we aim to examine to what extent the patterns of change in *be like* described in these North American varieties are also observed in England. Second, we aim to assess evidence for change in constraints on *be like* in real time using evidence from a trend study of usage and a controlled judgment study.

### **A Real-Time Study of Quotative Change in the U.K.**

#### *Data and Method*

To test Tagliamonte and D'Arcy's predictions about grammaticalization of *be like* we compare patterns of variation in quotative usage in two data sets, which we describe in turn below.

We gratefully acknowledge the assistance of Sali Tagliamonte and Rachel Hudson in providing us with data from the York storytelling corpus, first reported on in Tagliamonte and Hudson (1999). This data set is a corpus of one-on-one sociolinguistic interviews with 44 University of York undergraduates collected in the summer of 1996. The interviewers were fellow University of York undergraduates working under the direction of Sali Tagliamonte.

Tagliamonte and Hudson (1999) do not specify the ages of the participants in this corpus; however, in view of University of York undergraduate enrollment during this period, it is likely that all or nearly all of the subjects were between the ages of 18 and 22. All of the 44 subjects, who were evenly divided by sex, were native speakers of U.K. English. These data produced 397 tokens.<sup>2</sup>

For the purpose of inferring generational change in quotative use, we compare Tagliamonte and Hudson's 1996 data with a similarly constructed data set gathered in the spring of 2006. This data set consists of one-on-one interviews of 31 University of York undergraduates—14 women and 17 men—collected by the latter four authors. The interviews were conducted using a standard battery of questions intended to elicit maximally unselfconscious narratives, comparable to the data in Tagliamonte and Hudson's 1996 corpus. All the subjects are native speakers of U.K. English. From transcripts of these interviews all instances of quoted speech were extracted and coded following Tagliamonte and Hudson's (1999) procedure. These data yielded 955 tokens of quotatives.

To be able to fully compare the two data sets and to address the above issues to do with grammaticalization, we defined our token set slightly differently from some previous studies. First, tokens of quotative form *it's like*, as in (15), where the subject is impersonal, were excluded from the analysis.

(15) As soon as he came back in the room, it's like "Oh no!" (2006 data set)

Although these tokens are found in both data sets and at roughly equivalent rates (26 tokens, i.e. 5 percent of total in 1996; and 41, i.e. 4 percent of total in 2006), following Tagliamonte and D'Arcy's (2004:504) procedure we excluded *it's like* tokens as they have "an exceptional status." The occurrence of the pronoun *it*, for example, differentiates *be like*

from other quotative verbs which can appear only with subjects that are in some way animate. The construction is thus not fully variable.

Second, to facilitate comparison with Tagliamonte and D’Arcy’s (2004, 2007) results, we excluded tokens of non-lexicalized sounds. This removed 39 tokens in the 1996 data set and 80 in the 2006 data.

Third, because we are studying the effect of morphological tense on variation in quote introducers, we excluded non-tense-bearing forms including zero quotatives, infinitives and participles. This culling removed 90 tokens from the 1996 data set and 127 tokens from the 2006 data set. This protocol departs from previous studies but was required for the present data set in order to fit a coherent model (Tagliamonte & D’Arcy 2004, 2007; Tagliamonte & Hudson 1999). Removing these tokens allowed us to focus more closely the factors affecting *be like*.

Finally, our initial analysis of the use of quote introducers with progressives and participial adjuncts, as in (16) and (17), showed that this context was not fully variable.

(16) The cat is going “Meow.” (2006 data set)

(17) And he was standing there going “Ooh not again girls.”(2006 data set)

Recall that the use of *be like* in these contexts could be interpreted as a sign of further grammaticalization of the quotative system. In these environments, in which state predicates are canonically poor, *be like* was found to be strongly disfavored, although there were some signs of increase of use with *be like*. The 1996 data set contains 94 such examples (about 23 percent of the overall number of tokens), none of which occur with *be like*. In the 2006 data set, there are 105 such tokens (accounting for 12 percent of the overall number), two of which contain *be like*.<sup>3</sup> For these reasons, we treated this variation as categorical and

excluded all tokens in these environments from the analysis.

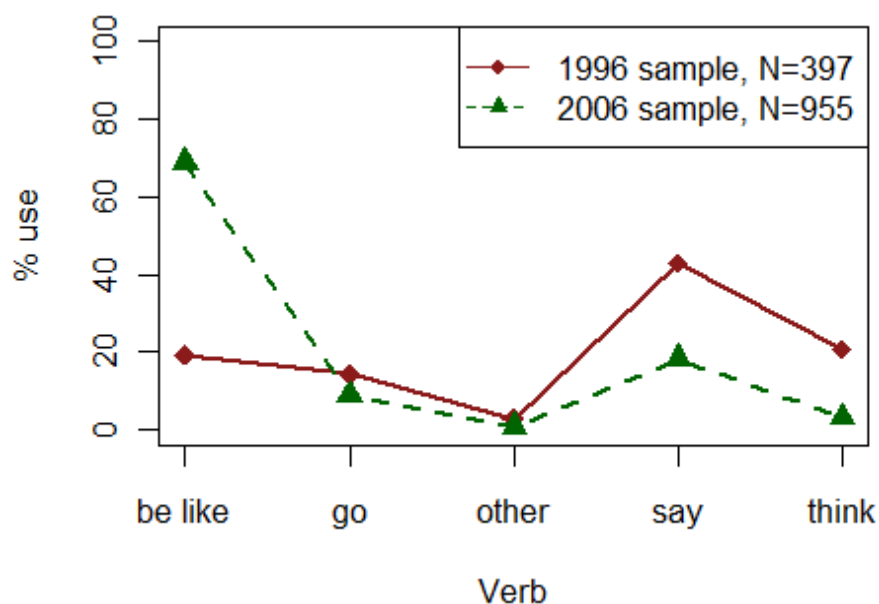
The near categorical absence of *be like* in these environments suggests an important limit on the extent to which *be like* has come to be reanalyzed as a quotative on a par with accomplishment verbs like *say* and *go*. The presence in our data of these two tokens of *be like* progressives demonstrates that for some speakers, at least, *be like* may indeed behave as a true event (non-state) predicate.<sup>4</sup> Nevertheless, the relative scarcity of such tokens indicates that *be like*'s conservative competitors remain preferred in usage in unambiguously eventive contexts in this sample.<sup>5</sup> We return to these semantic considerations later in this paper.

In order for our analysis to be as comparable as possible to previous multivariate results, we follow an approach used first by Tagliamonte and Hudson (1999) and subsequently elsewhere (Tagliamonte & D'Arcy 2004, 2007; Buchstaller & D'Arcy 2009) whereby all of *be like*'s main competitors, including *go*, *say*, and *think*, are included in the envelope of variation. *Think* and *say*, are not, of course, semantically equivalent since the former tends to introduce reported thought quotes and the latter direct speech. We note, however, that the adoption of this approach has no consequence for our main measures of interest, namely the interaction between age and the internal and external constraints introduced above.

### *Results*

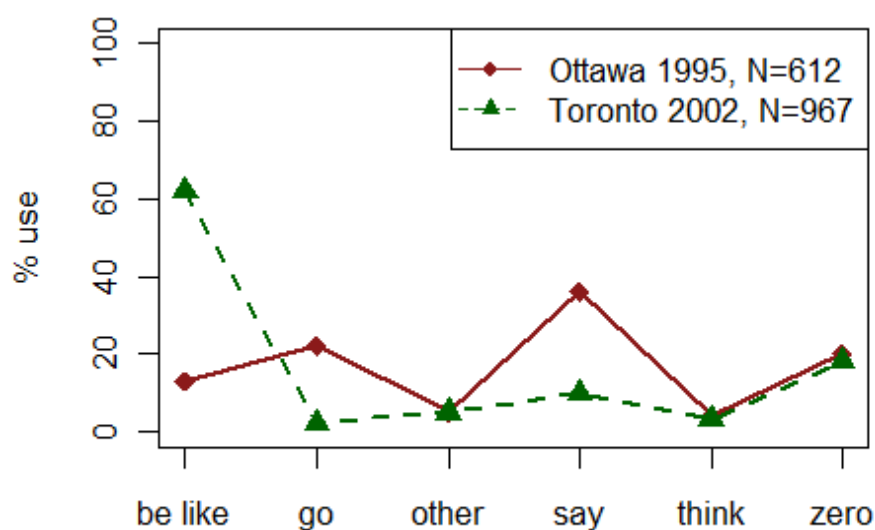
A comparison of these two data sets suggests a sharp increase in *be like* usage over the 10 years between 1996 and 2006. Among York undergraduates, *be like* has become the most frequently used quotative, increasing from 19 percent in 1996 to 68 percent in 2006. This increase comes at the expense of *be like*'s main competitors—*say*, *think*, and, to some extent, *go*—, all of which show lower rates of use in the 2006 sample than in the 1996 data. Use of other quotatives—*shout*, etc.—is marginal in both samples. Figure 1 provides an overall distribution of five of the most frequent quotatives in our two data sets.

**Figure 1:** Overall distribution of variants for two data sets



A similarly vertiginous increase in *be like* use is also reported in a study of Canadian youth by Tagliamonte and D’Arcy (2004). This study compares patterns of quotative use in two corpora of sociolinguistic interviews—one consisting of data from University of Ottawa students in 1995, and a second from Toronto youth aged 10-19, collected in 2002 and 2003. Like the present York data, Tagliamonte and D’Arcy’s Canadian data show that *be like* has become the most frequent quotative used in narratives among young people, rising from 13 percent in the 1995 Ottawa data to 63 percent in the Toronto data for 17-19 year olds (the age group closest to the Ottawa University students). We reproduce Tagliamonte and D’Arcy’s results in Figure 2. The present data, then, suggest that *be like* is diffusing into the speech of British youth with a pace similar to that described in the literature on North American Englishes.

**Figure 2:** Tagliamonte and D’Arcy’s real time results from Canada.



A further goal of this paper is to test change in real time in the effect of the well-studied social and linguistic constraints on *be like* use described earlier. To this end, we fit a logistic regression model using both data sets and tested for significant interactions between our sample factor (with levels 1996 sample and 2006 sample) and our four social and linguistic factors: speaker sex, subject person, quote content, and morphological tense. The model is summarized in Table 3, which shows the contribution of different factors to *be like* usage. The first column lists the factors in the model. The second column gives log odds ratios for the treatment level given in brackets,<sup>6</sup> and the next three columns provide the standard error, z-score, and p-values for these coefficients.

**TABLE 3**

Contribution of factors to *be like* usage

Factor	Coefficient	St. Err.	Wald Z	P
Intercept	-0.5547	0.1773	-3.13	0.0018
Sample (2006)	1.7704	0.1905	9.29	0.0000
Sex (male)	-1.5562	0.2879	-5.40	0.0000
Person (third)	-0.8101	0.1293	-6.26	0.0000
Tense (present)	0.8448	0.1718	4.92	0.0000
Content	--	--	--	ns
Sample :sex*	1.1479	0.3224	3.56	0.0004

Sample :person	--	--	--	ns
Sample :content	--	--	--	ns
Sample :tense	--	--	--	ns

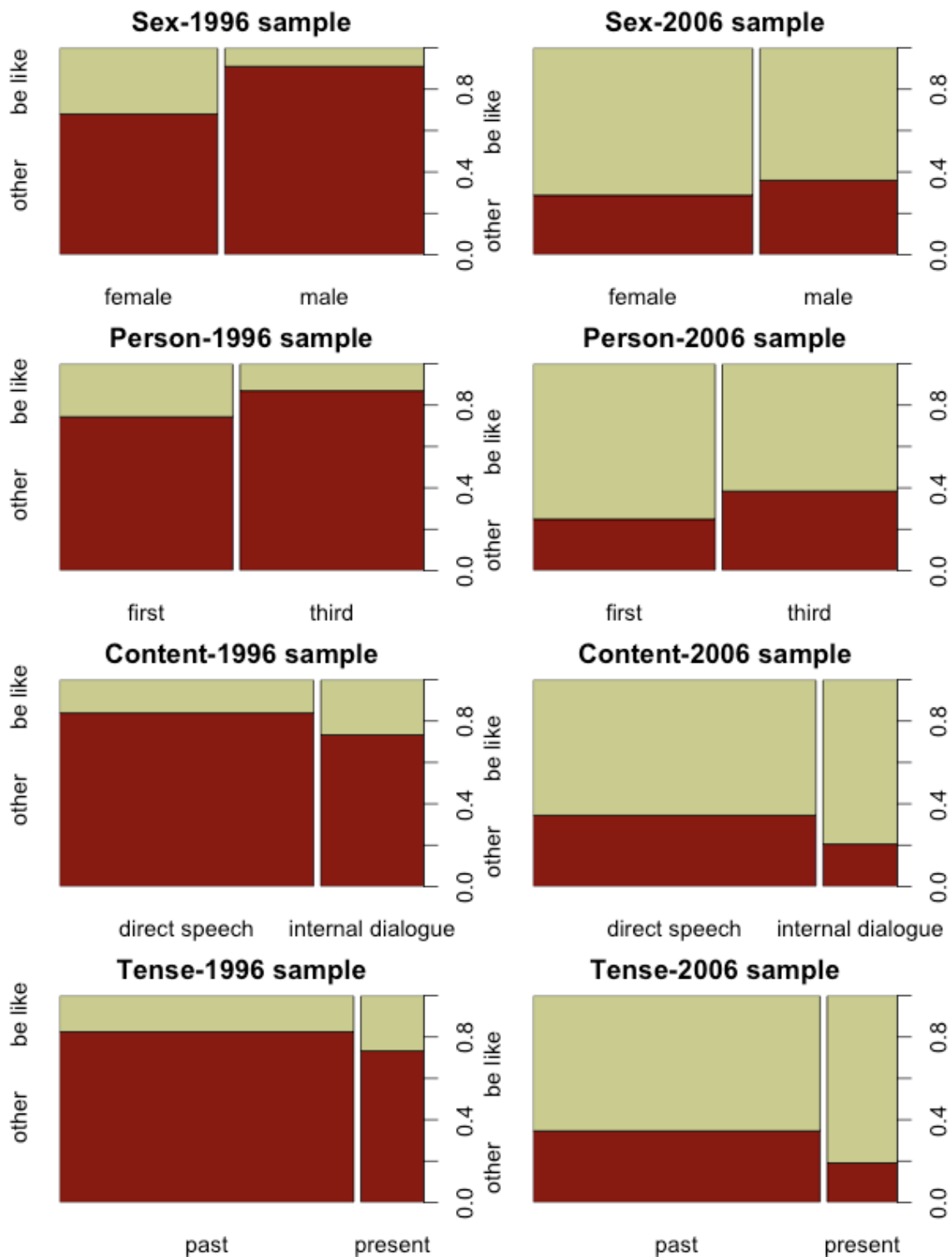
\* With treatment levels 2006 sample and Male.

We illustrate these effects in Figure 3, which shows proportional use of *be like* versus other quotatives by condition for the two samples.

The analysis summarized in Table 3 reveals a significant main effect for sample. As also illustrated in Figures 1 and 3, *be like* is much more frequent in the 2006 corpus than the 1996 corpus. In addition, Table 3 shows a significant main effect for sex, with women favoring *be like*, and a significant interaction between sex and sample. Figure 3 demonstrates that the effect of speaker sex in the 1996 sample is much greater than in the 2006, suggesting an attenuation of the sex effect over time.

These data are surprising from the perspective of Tagliamonte and D’Arcy’s (2004, 2007) studies of quotative use in Toronto, Canada, where the youngest speakers, with the highest rates of *be like* use, showed the strongest sex effects. Older speakers in the sample, who Tagliamonte and D’Arcy suggest were among the *be like* innovators in the community a generation earlier, used *be like* relatively little and showed a much weaker sex differentiation. Based on these data, Tagliamonte and D’Arcy propose that sex differentiation in *be like* use has emerged in Toronto as it continues to diffuse into local speech.

**Figure 3:** Proportional use of *be like* vs. other quotative verbs by speaker sex and context



The present results, in contrast, suggest a mild form of neutralization of the speaker sex effect (Ferrara & Bell 1995). These findings, alongside Tagliamonte and D'Arcy's (2004, 2007) results, imply that the diachronic development of the sex effect in *be like* usage is



variable across communities. One possible explanation of this U.K.-Canadian difference in sex effects is that it represents different stages of change for *be like*. Much of the diachronic literature has suggested that sex effects vary across the trajectory of change, with strong sex associations appearing at early stages of change and a weakening effect appearing at later stages (Labov 2001:308; Tagliamonte & D'Arcy 2009). From this perspective, one possibility is that the diffusion of *be like* is slowing in the U.K. but not Canada, which would explain why the effect of speaker sex is becoming neutralized.<sup>7</sup> We are aware of no independent evidence in support of this supposition, however.

The divergent sex effects in the U.K. and Canada are nevertheless in keeping with some matched guise evidence suggesting a much stronger gender association for *be like* in North America than in Britain. In particular, in Dailey-O'Cain's (2000) U.S. study, 80 percent (24/30) of subjects identified *be like* with women. In a similarly constructed study in the U.K., Buchstaller (2006) found that only 34 percent (65/191) of subjects associated *be like* with women. Our data lend further support to Buchstaller's suggestion of cross-societal differences in the social meaning of using *be like*. As *be like* continues to diffuse globally, the social meaning associated with its use does not necessarily diffuse along with the surface form. Rather, individual communities adapt the innovation in the context of local social and economic conditions and local symbolism (Eckert 2000).

The quote content data in Table 3 are also unexpected from the perspective of much previous literature in that the analysis returned no significant main effect or interactions for quote content. Figure 3 shows that, in both of the above samples, *be like* is used more frequently in reported thought contexts than with direct speech, but this difference does not contribute significantly to the model. This result, again, contrasts with Tagliamonte and D'Arcy's Canadian findings (2004, 2007) which propose an expansion of *be like* into direct speech contexts over time. In particular, Tagliamonte and D'Arcy report that the youngest

and most advanced set of *be like* users—the 17-19 year olds—favor *be like* in direct speech contexts and disfavor it in internal dialogue. The present data provide no evidence of a change in the effect of the reported thought/direct speech contrast. The fact that the analysis summarized in Table 3 and Figure 3 returned no main effect for quote content suggests the possibility that the effect of quote content weakened in England at an earlier stage relative to overall usage of *be like* than in Canada. That is, in Canada, this effect weakens substantially only among cohorts for which *be like* is the majority variant in spoken interview-style data. In the U.K. sample, this effect is fairly weak even among the earlier users for whom *be like* is a minority variant (see Figures 1 and 2). We return to these issues later.

The effect of subject person in Table 3 and Figure 3 is negligible. The analysis returned a significant main effect for subject person, with first-person subjects favoring *be like*. No significant interaction with sample was returned, indicating no evidence of inconstancy in the effect of subject person between the two samples, as illustrated in Figure 3. This result is expected from the perspective of most previous studies, which suggest a remarkable consistency of the subject person effect across contexts (Buchstaller & D’Arcy 2009; Tagliamonte & D’Arcy 2004, 2007; Tagliamonte & Hudson 1999).

The tense results in Table 3 are also in keeping with results in previous literature. Table 3 and Figure 3 show that morphological present tense favors *be like* while past tense favors other quotatives (Singler 2001; Tagliamonte & D’Arcy 2007). No significant interaction with sample year was returned, suggesting no inconstancy in the effect of tense over time.

From the perspective of the literature reviewed above, the most important outcome in these data is the difference between speaker sex and the three linguistic factors—tense, subject person and quote content—in terms of their interaction with sample group. Again, the data show a weakening of the sex effect from 1996 to 2006, but a constancy of the three linguistic factors. The mutability of the speaker sex effect in itself is unremarkable in light of

much previous literature on *be like* and other phenomena showing that the way linguistic variation indexes social meaning often changes over time in a community as the variable undergoes social reanalysis and constant symbolic re-appropriation (Dyer 2002; Eckert 2000, 2008; Tagliamonte & D'Arcy 2004). The fact that the sex effect in our trend study has changed in the direction opposite to that indicated in Tagliamonte and D'Arcy's studies is in keeping with results from Buchstaller and D'Arcy's (2009) multi-community comparison. This suggests that, as *be like* has diffused geographically, social constraints on quotative variation often change from community to community, as patterns of linguistic variation map onto local social and stylistic differences in community-specific ways.

The constancy of the linguistic factors contrasts with findings from much of the literature implying that there is a change in the effect of some of these constraints (Ferrara & Bell 1995; Tagliamonte and D'Arcy 2004, 2007). The gap between these findings and previous corpus results from other locales therefore supports skepticism with regard to universals of grammaticalization of *be like* that dictate change in linguistic effects in *be like*, particularly quote content. We consider these issues further in light of judgment data presented in the following discussion.

### **Direct Speech and Non-Speech Interpretations of *Be Like* Quotatives in Judgment Data**

The second set of data for this study comes from a judgment experiment focusing on age effects on direct speech and reported thought interpretations of *be like* and *say* quotatives. As discussed above, early work on *be like* described it not as an introducer of direct speech, but rather exclusively as an introducer of reported thought. This variant of *be like* may have emerged as a reanalysis of descriptions of states of individuals in sequences of *be* + focuser/discourse marker *like* + predicate adjectives or non-lexicalized sounds (Butters 1982; Tannen 1986) (see examples in (4) and (5)).

Much subsequent corpus-based work on *be like* however has reported that quotes introduced by *be like* could be used to describe not just states of individuals as in (18a), but also saying eventualities as in (18b).

- (18) Aaron was like “Ok, fine.”
- a. ‘Aaron thought/felt like saying “Ok, fine.”’
  - b. ‘Aaron said “Ok, fine.”’

A disadvantage of usage corpora for analyzing semantic variation of this kind is that the intended reading can be difficult to discern from the context. To complement the corpus-based findings discussed above, we report on a controlled judgment experiment intended to examine cross-speaker differences in the availability of *be like* quotatives in contexts biasing these different readings. If younger speakers are coming to reanalyze *be like* quotative predicates as descriptions of speech events rather than reported thought, then we expect age to interact with the reported thought/direct speech difference in judgments of quotative sentences. The following discussion describes an experiment designed to test this claim. Acceptability judgments are a different kind of linguistic performance from speech in a sociolinguistic interview context, which raises the question of the comparability of judgment data and speech corpus data for the present purposes. Several recent papers in fact have suggested that judgments of acceptability are sensitive to and in fact closely mirror the relative probabilities of semantically equivalent competing forms in production (Bresnan 2007; Bader & Häussler 2010a, 2010b; Melnick et al. 2011). We assume, based on these studies, that acceptability judgments are indeed a suitable kind of data for the purpose of inferring cross-speaker differences in linguistic effects on quotative variation.

### *Data and Method*

The participants were 121 self-described native speakers of American English aged 18-73 (M=31.3, SD=11.6)—71 women and 50 men.<sup>8</sup> All had at least some university education, and in this respect our sample is similar to those reported on in several previous studies on *be like* change (Ferrara & Bell 1995; Dailey-O’Cain 2000; Tagliamonte & Hudson 1999).

Previous literature on *be like* has indicated no strong regional effects on *be like* usage within the U.S. and no effort was made to stratify the sample by region. No regional effects were observed in the analysis. Participants were recruited online through the contacts of the researchers and were not paid for their participation.

The experiment compares scores for *be like* and *say* sentences in six environments. A first, baseline context was created with no stativity/eventivity bias, as in (17). Four additional contexts—progressives, imperatives, *force...to* complements, and pseudoclefts with *do*—were used as ways of biasing eventive readings; all of these are contexts in which eventive predicates are fine, but true states are poor (Dowty 1979). We illustrate this contrast in (19)-(22), which compare stative *have \$100* with eventive *spend \$100* in each environment.

- (19) She was \*having \$100/spending \$100. (progressives)
- (20) Just \*have \$100/spend \$100. (imperatives)
- (21) Tim forced him to \*have \$100/spend \$100. (*force...to*)
- (22) What she needs to do is \*have \$100/spend \$100. (*do* pseudoclefts)

In the following experiment, we use these environments to compare acceptability of eventive, direct-speech readings of *be like* and *say*, as illustrated in (23)-(26).

- (23) She was being like/saying, “They’re coming tomorrow at 11:00” (progressives)

- (24) Just be like/say, “They won’t ever do it.” (imperatives)
- (25) Tim forced him to say/be like, “Fine, I’ll do it next week.” (*force...to*)
- (26) What she needs to do is say/be like, “John already quit.” (*do pseudoclefts*)

The final environment biased non-speech *be like* readings using *for* adverbials. As illustrated in (27), temporal *for* phrases are fine with atelic predicates in simple tenses but poor with eventives (Dowty 1979).

- (27) For an hour, Mark had \$100/\*spent \$100. (*for adverbials*)

We use such contexts to diagnose the availability of stative, non-speech interpretations of *be like* and *say* quotative predicates, as in (28).

- (28) For an hour, Mark was like/said, “Let’s go to McDonald’s.” (*for adverbials*)

Two lexicalizations were created for each environment, each assigned either to a *be like* or *say* condition yielding two test sets. Each participant therefore saw each condition once. Subjects were randomly assigned to test sets, and a unique random order of the 12 test sentences and 18 fillers was created for each subject. A list of the experimental sentences is provided in Appendix 1.

The data were gathered through a self-paced online magnitude estimation procedure using WebExp2 software (Mayo et al. 2008) in the summer of 2009. In syntactic magnitude estimation experiments, subjects judge stimulus sentences not on an abstract n-point scale but rather in relation to a positive numerical score arbitrarily assigned to a benchmark (“modulus”) sentence (Bard et al. 1996; Keller & Sorace 2003). If the stimulus sentence is

judged to be twice as acceptable as the benchmark sentence, the participant gives it twice the benchmark score; if it is half as acceptable, half the benchmark score, and so on. In the present experiment, the benchmark sentence used was that in (29), which native speakers of English typically find to be of intermediate well-formedness.

(29) I wouldn't give to the boy the difficult puzzle.

After they had given consent to participate, subjects were asked to provide some background information, including age, sex, highest level of education completed, and hometown. Subjects were then introduced to the magnitude estimation procedure, and then given two sets of slides providing practice in applying this technique. In the first set, subjects used magnitude estimation to measure lengths of lines; the second set provided sample sentences to judge. The experimental phase followed, which subjects typically completed in five to ten minutes.

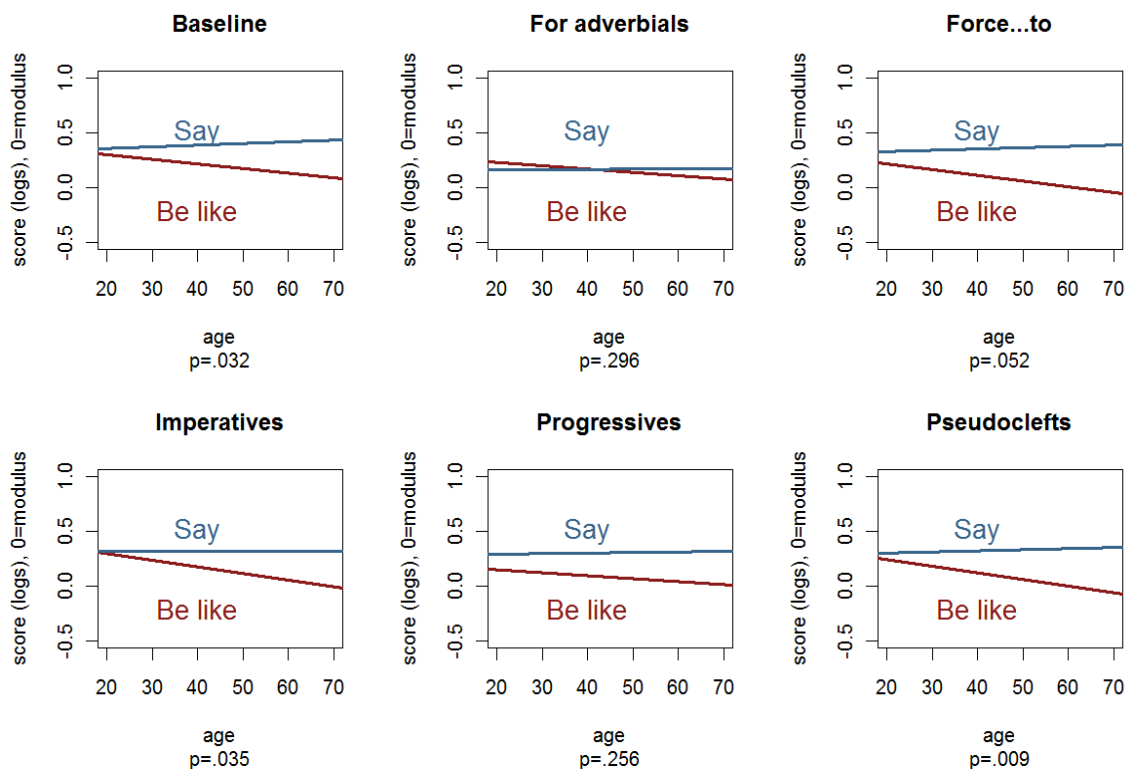
Following Bard et al.'s (1996) procedure, raw scores were normalized by dividing them by the benchmark score. The base-10 logarithms of these scores were then taken in order to make data normally distributed and suitable for parametric tests. In the following discussion, we report these normalized, log-transformed values.

### *Results*

To examine the effect of speaker age on acceptability scores, we fit mixed-effect linear models for each condition using the lme4 package for R (Bates & Maechler 2010; R Development Core Team 2008). The dependent variable was the log-transformed values for each condition, with age and verb as fixed effects and subject and item as random effects. Following Baayen's (2008) procedure, p-values were simulated by Markov chain Monte

Carlo (MCMC) sampling (10,000 samples) using the LanguageR package for R (Baayen 2006, 2010). To examine cross-generational difference in acceptance of *be like* in each of these environments, we focus not on the effect of age but rather the age-verb (*be like* vs. *say*) interaction. We choose this measure in order to account for a possible age effect in preferences toward direct speech versus reported speech. The results are summarized in Figure 4, which plots *say* and *be like* scores by subject age for each condition and reports a p-value for the verb:age interaction variable. Zero on the y-axis corresponds to the subject's value for the benchmark sentence in (28). Values above zero, therefore, reflect a judgment for the stimulus sentence better than that for (29), and scores below zero reflect a judgment below that for (29).

**Figure 4:** *Say* and *be like* scores by age for six conditions.



The plots in Figure 4 show, unsurprisingly, that the *say-be like* acceptability gap is inversely correlated with age. Subjects under 30 years of age in the sample accept *be like* in



each of these six conditions, roughly on a par with *say*, indicating that *be like* is natural for many young speakers in contexts biasing both eventive and stative interpretations of *be like*. Nevertheless, the interaction between age and verb reaches significance at  $\alpha=.05$  only for three of these environments: the baseline context, pseudoclefts, and imperatives. The interaction for *force...to* complements is suggestive at  $p=.052$ . For *for*-adverbials there is no interaction between age and verb, and in fact no main effect for verb. These judgment data therefore align only partially with corpus data suggesting diffusion of *be like* in direct speech and non-speech contexts. The absence of more consistent age effects in these data may be partially attributable to the fact that our sample is relatively youthful, with a mean age of 31.1.

More directly relevant to the issues considered here is the fact that the regression lines for the *be like* conditions in Figure 4 are all roughly parallel. To examine possible interaction between age and the speech/non-speech contrast, we fit four separate mixed effect linear models with scores for *for*-adverbials and each of the four speech event-biased conditions as levels in a fixed factor; age was an additional fixed effect with random effects speaker and item. The analyses returned no significant interaction at  $\alpha=.05$  for any of the four comparisons, suggesting no evidence of an age difference in the effect of the direct speech/reported thought contrast. (For *for* adverbials vs. *force...to* as the eventive/stative comparison,  $p=.542$ ; *for* adverbials vs. pseudoclefts  $p=.316$ ; *for* adverbials vs. progressives  $p=.930$ ; and *for* adverbials vs. imperatives,  $p=.317$ .) These judgment results from U.S. English speakers align with the corpus results presented earlier in that they suggest no change in the effect of the reported thought/direct speech contrast over time, which diverges from the findings reported in Ferrara and Bell (1995) and Tagliamonte and D'Arcy (2004, 2007). The contrast between our findings and previous corpus findings from other locales again supports skepticism toward strong universal pathways of grammaticalization of *be like*.

## Conclusion

Results from two studies—one a usage study and the other a controlled judgment study—suggest no interaction between age and the typical linguistic effects described in foundational work by Ferrara and Bell (1995) and Tagliamonte and Hudson (1999). Most importantly, results from our studies fail to support an age difference in the effect of the direct speech/reported thought contrast as reported in Tagliamonte and D’Arcy (2004, 2007). From the perspective of Tagliamonte and D’Arcy’s proposals and likeminded work suggesting rigid grammaticalization pathways in syntactic change (Jespersen 1917; Roberts and Roussou 1999; Saxena 1995), one possible interpretation of the above results, alluded to earlier, is that the Toronto dialect sits at a different point in the trajectory of grammaticalization from the varieties represented by our U.S. and U.K. samples. Again, the multivariate analysis summarized in the section examining the longitudinal data turned no significant main effect or interactions for quote content. This result suggests the possibility that the leveling of the quote content effect hypothesized by Tagliamonte and D’Arcy came about at an earlier stage than in Canada relative to overall usage of *be like*. This would mean that the reanalysis responsible for the change in contextual effects on *be like* is independent of overall frequency of usage of *be like*, a plausible assumption.

While the constancy in linguistic effects in the two data sets reported on here are surprising from the perspective of some previous corpus findings, they are expected from the perspective of findings in the quantitative diachronic syntax literature. Much previous quantitative work on historical corpora has shown that for any single abstract process of syntactic change, contextual effects are typically constant over time—a phenomenon known as the *constant rate effect* (Freuhwald et al. 2009; Kroch 1989, 1994, 2000; Pintzuk 1991; Santorini 1992).

In particular, Kroch (1989, 2000) attributes this constancy to individuals' language-independent faculty for tracking frequencies of experienced events. As learners acquire and increment new forms, they learn from input sources the relative propensities of use of variants in different contexts, with the consequence that contextual effects are propagated across generations of speakers, all other things being equal. Occasionally, linguistic factors can come to interact with social factors in new ways, which may have the effect of changing the effects across time, but this is the exception rather than the rule, to judge from the published literature (Kroch 1989, 2000). Kroch and colleagues' model and the present results lead us to anticipate that the typical pattern of evolution for *be like* will not involve changing linguistic constraints as in Tagliamonte and Darcy's (2004, 2007) findings, but rather will be characterized by consistency in linguistic effects on *be like*. Future work might usefully address this possibility.

### **Acknowledgments**

We are grateful to the subjects of this study and to Sali Tagliamonte and Rachel Hudson for sharing their 1996 data set with us. Thanks also to Helen Lawrence and colleagues at the University of York for help in contacting potential informants. Thanks to Alex D'Arcy, Tony Kroch, Sali Tagliamonte and audiences at NWAV, DiGS, LingEvid 2010, and University of Glasgow for comments on some of the material presented here. This research is supported by ESRC grant number 061-25-0033. We are responsible for all errors.

### **Notes**

1. Other well-known properties distinguishing state and activity, including the fact that the former but not the latter are typically poor in imperatives and pseudoclefts with *do*, are not useful for our purposes since these contexts arise very rarely in corpus data.

2. This is a slightly smaller N from that reported in Tagliamonte and Hudson (1999)—665. We have excluded some of Tagliamonte and Hudson’s tokens for reasons detailed in the methodology.
3. These tokens are given in (i) and (ii).
  - (i) Driving through there being like “Argh.”
  - (ii) And just being like “Suzie can you please get my flip flops for me.”
4. In an appendix Tagliamonte and D’Arcy (2007:217) also report a handful of *be like* tokens in progressive contexts and with present participles.
5. A similar issue concerns the behavior of quotatives with recipient arguments. Conservative quotative variants in English all allow for argument structures with a recipient argument, either in a prepositional dative construction (with rightward shift of the quoted material) (iii) or double object construction (iv).
  - (iii) He’d go to her “Just listen for goodness sake”
  - (iv) I asked her “Are you crazy?”

We have no instances of *be like* in our relatively small set of tokens in prepositional dative and double object constructions (~20); all of these are with *go* or *say*. Indeed, sentences such as (v) sound fairly unnatural to native speakers we have consulted.

  - (v) ?She’d be like to them “You’ve got to shut up.”

The hypothesis of grammaticalization of *be like* suggests the possibility of the eventual emergence of such sentences. Indeed, quotative *go* appears to have undergone a similar process of change. That is, quotative *go*—which derives from verb of motion *go*—now happily takes a *to* + recipient PP as in (iii).
6. This measure, akin to a factor weight in Varbrul analyses, represents the effect of the contrast with the competing baseline condition (not shown) with a value of zero. In

Table 3, for example, third-person contexts have a negative effect relative to first-person contexts (-0.8101) on the probability of *be like* usage.

7. We owe this suggestion to Alex D'Arcy.
8. We chose American English speakers for convenience. We note that the goal of this experiment is not to inquire into the relationship between judgment data and variation in usage corpora *per se*, but rather to assess via a different kind of data the possible interaction between speaker age and the direct speech-reported thought contrast. The use of a different population for this experiment is therefore innocuous for the purposes of the claims addressed here.

## References

- Bader, Markus & Jana Häussler. 2010a. Toward a model of grammaticality judgments. *Journal of Linguistics* 46. 273-330.
- Bader, Markus & Jana Häussler. 2010b. The primacy of grammaticality. Ms. University of Konstanz.
- Bard, Ellen, Dan Robertson & Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1). 32-68.
- Baayen, R. Harald. 2006. *Analyzing linguistic data*. Cambridge: Cambridge University Press.
- Baayen, R. Harald. 2010. LanguageR library for R. <http://cran.r-project.org/>. Accessed September 10 2011.
- Bates, Doug & Martin Maechler. 2010. Lme4 package for R. <http://cran.r-project.org/>. Accessed September 10 2011.
- Blyth, Carl, Sigrid Recktenwald & Jenny Wang. 1990. I'm like, "say what?!": A new quotative in American oral narrative. *American Speech* 65. 215-227.
- Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the

- English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base*, 77-96. Berlin: Mouton de Gruyter.
- Buchstaller, Isabelle. 2006a. Diagnostics of age-graded linguistic behaviour: The case of the quotative system. *Journal of Sociolinguistics* 10(1). 3-30.
- Buchstaller, Isabelle. 2006b. Social stereotypes, personality traits and regional perception displaced: Attitudes towards the 'new' quotatives in the UK. *Journal of Sociolinguistics* 10(3). 362-381.
- Buchstaller, Isabelle & Alex D'Arcy. 2009. Localized globalization: A multi-local investigation of quotative *like*. *Journal of Sociolinguistics* 13(3). 291-331.
- Butters, Ronald. 1982. Editor's note [on 'be like']. *American Speech* 57. 149.
- Cukor-Avila, Patricia. 2002. She say, She go, She be like: Verbs of quotation over time in African American Vernacular English. *American Speech* 77. 3-31.
- Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of Sociolinguistics* 4. 60-80.
- Dowty, David R. 1979. *Word meaning and Montague grammar*. Dordrecht: D. Reidel.
- Dyer, Judy. 2002. 'We all speak the same round here': Dialect levelling in a Scottish-English community. *Journal of Sociolinguistics* 6. 99-116.
- Eckert, Penelope. 2000. *Linguistic variation as social practice: The linguistic construction of social meaning in Belten High*. Oxford: Blackwell.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12. 453-476.
- Ferrara, Kathleen & Barbara Bell. 1995. Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of *be like*. *American Speech* 70. 265-289.
- Freuhwald, Josef, Jonathan Gress-Wright & Joel Wallenberg. 2009. Phonological rule change: The constant rate effect. In Seda Kan, Claire Moore-Cantwell and Robert

- Staub (eds.), *NELS 40: Proceedings of the 40<sup>th</sup> meeting of North East Linguistic Society*. Amherst: GLSA.
- Jespersen, Otto. 1917. *Negation in English and other languages*. (Historiskfilologiske Meddelelser I.5.) Kopenhagen: A.F. Høst.
- Keller, Frank & Antonella Sorace. 2003. Gradient auxiliary selection and impersonal passivization in German: An experimental investigation. *Journal of Linguistics* 39. 57-108.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199-244.
- Kroch, Anthony. 1994. Morphosyntactic variation. In Katharine Beals (ed.), *Proceedings of the 30th Annual Meeting of the Chicago Linguistics Society*, 180-201. Chicago: Chicago Linguistic Society.
- Kroch, Anthony. 2001. Syntactic change. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 699-730. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Macaulay, Ronald. 2001. You're like 'why not?'. The quotative expression of Glasgow adolescents. *Journal of Sociolinguistics* 5. 3-21.
- Mayo, Neil, Martin Corley & Frank Keller. 2008. WebExp2, [www.webexp.info/](http://www.webexp.info/).
- Melnick, Robin, T., Florian Jaeger & Thomas Wasow. 2011. Speakers employ fine-grained probabilistic knowledge. Paper presented at the 85<sup>th</sup> Annual Meeting of the Linguistic Society of America, Pittsburgh, PA.
- Pintzuk, Susan. 1991. Phrase structures in competition: Variation and change in Old English word order. Philadelphia, PA: University of Pennsylvania Ph.D. dissertation.
- R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing,

- Roberts, Ian & Anna Roussou. 1999. A formal approach to ‘grammaticalization’. *Linguistics* 37(6). 1011-1041.
- Romaine, Suzanne & Deborah Lange. 1991. The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66. 227–279.
- Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language & Linguistic Theory* 10. 595–640.
- Saxena, Anju. 1995. Unidirectional grammaticalization: diachronic and cross-linguistic evidence. *Sprachtypologie und Universalienforschung* 48(4). 350–372.
- Singler, John. 2001. Why you can’t do a Varbrul study of quotatives and what such a study can show us. *University of Pennsylvania Working Papers in Linguistics* 7. 257–278.
- Tagliamonte, Sali A. & Alex D’Arcy. 2004. “He’s like, she’s like”: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8. 493–514.
- Tagliamonte, Sali A & Alex D’Arcy. 2007. Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* 19(2). 1-19.
- Tagliamonte, Sali A. & Rachel Hudson. 1999. *Be like* et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3. 147–172.
- Tannen, Deborah. 1986. Introducing constructed dialogue in Greek and American conversational and literary narrative. In Florian Coulmas (ed.), *Direct and indirect speech*, 311-332. Amsterdam: Mouton de Gruyter.



## **Appendix 1: Experimental Sentences**

### ***Baseline:***

Jenny was like/said, "They're coming at 11:00.

Sam was like/said, "It's tomorrow, Mom."

### ***Pseudoclefts with do:***

What she needs to do is be like/say, "John already quit."

What he should do is be like/say, "Come by on Tuesday."

### ***Imperatives:***

Just be like/say, "They won't ever do it."

Just be like/say, "You can come if you want."

### ***Force..to.***

Tim forced him to be like/say, "Fine, I'll do it next week."

Maria forced her to be like/say, "Yes, I own a guitar."

### ***Progressives.***

Janet was being like/saying, "That's never going to work."

Emma was being like/saying, "He didn't believe it."

### ***For adverbials***

For an hour, Mark was like/said, "Let's go to McDonald's."

For almost 45 minutes, Tammy was like/said, "I'm tired of your criticism."